Analysis Plan

Project Name: Reducing Inappropriate Prescribing and Co-Prescribing Behaviors: A Retrospective Analysis

Project Code: 2002

Date Finalized: 4/28/2022

Project Description

Improper prescribing of opioids represents a major public health issue, and prescribed opioids are involved in nearly one third of all opioid overdose deaths.[1] What's more, efforts to mitigate improper prescribing are timely given Provisions 6065 and 6052 of the SUPPORT Act, which mandate sending peer comparisons letters to high prescribers of opioids.[2]

The Office of Evaluation Sciences (OES) is collaborating with the Center for Program Integrity (CPI) at the Centers for Medicare & Medicaid Services (CMS) to better understand the efficacy of interventions aimed at addressing this key issue. CPI oversees efforts to safeguard the Medicare and Medicaid programs from fraud, waste and abuse. CPI mailed peer comparisons letters to high co-prescribers of opioids and benzodiazepines in January 2020. Other similar letters were sent in January 2021 and January 2022; and previously in May 2019, letters were also sent to high prescribers of opioids. The aim of this collaboration is to provide recommendations on how to conduct a retrospective analysis of the impacts of letters sent in 2020 on prescribing behavior, and to inform the CMS response to Provisions 6065 and 6052 of the SUPPORT Act.

Data and Data Structure

This section describes variables that we recommend be analyzed, as well as changes that be made to the raw data with respect to data structure and variables.

Data Source(s):

Data would come from the CMS Integrated Data Repository (IDR), which tracks Medicare beneficiary enrollment healthcare utilization. The data span 10+ years, through the present. Prescriber data from July 1 2018-June 30 2019 were used to assess outlier status (defined as high volume prescribers being in the top 10% of co-prescribing opioids and benzodiazepines) for letters sent in 2020. We recommend that the team conducting the analysis examine prescribing behavior prior to letters being mailed, and focus on outcomes beginning one day after the letters were mailed through to 1 year, using baseline data from the prior year as controls.

[1] https://www.cdc.gov/drugoverdose/maps/rxrate-maps.html;
[2] https://www.govinfo.gov/content/pkg/PLAW-115publ271/html/PLAW-115publ271.htm

Key variables to be included would be drawn from the Medicare Part D Prescription Drug Event (PDE) data. These variables are used in the SAS code which extracts data from the IDR (see Appendix for details).

Outcome Variables to Be Analyzed:

The primary outcomes of interest are:

1. The **total MME (defined as average daily MME x days of co-prescriptions) among patients co-prescribed opioids and benzodiazepines** for the January 2020 letters and
2. The **percent of beneficiaries with overlapping prescriptions**.

Details on how these values are calculated in the IDR are available from the Department of Health & Human Services Office of Inspector General "Toolkit: Using Data Analysis To Calculate Opioid Levels and Identify Patients At Risk of Misuse or Overdose," available here.

Transformations of Variables:

We recommend computing total MME for each patient by summing the MME on each of the patient's opioid fills. In order to obtain MME levels per prescriber in our final prescriber-month level dataset, we recommend obtaining the average MME prescribed per month. MME is calculated by multiplying the milligrams of the active ingredient in the fill by a morphine equivalency conversion factor.

Imported Variables:

Opioids and benzodiazepines: In order to obtain the list of specific opioids and benzodiazepines used by prescribers in the PDE, we suggest referring to the subset of NDC codes that CMS initially used to classify prescribers as eligible or not for the letters. This same list may identify which drug codes will be relevant when measuring outcomes.

Outlier prescribers: There are 689 prescribers in the population who received a peer comparison letter.[3] The prescribers who were sent letters in 2020 were the top 10 percent of co-prescribers according to the average Morphine Milligram Equivalent (MME) per day of opioids prescribed to Medicare beneficiaries. The prescribers were selected from those who prescribed an opioid along with a benzodiazepine for at least 30 consecutive days to five or more Medicare patients in the last 12 months in their specialty and/or state. A detailed methodology of how the outlier group was selected is described in in CMS's "Methodology for Comparative Analysis: Co-Prescribing Patterns for Benzodiazepines with Opioids and Average Morphine Milligram Equivalents (MME)," available here. There is a natural control population of prescribers between the 80th and 90th percentile of MME per day and percent of overlapping prescriptions according to baseline

[3] An additional 406 prescribers received letters in January 2021 (of this sample, 218 also received similar letters in 2020). 322 prescribers received letters in January 2022.

prescribing behavior. These prescribers did not receive a letter and are identified in the IDR by CPI.

We do not anticipate requiring other imported variables or data.

Transformations of Data Structure:

We recommend that the analysis be conducted at the prescriber-month level such that the final dataset will have one observation per prescriber per month. Datasets provided for different time periods will be merged using prescriber NPI as the unique identifier.

Data Exclusion:

No observations should be excluded from the analysis.

Treatment of Missing Data:

Because the IDR data were used to identify the outlier prescribers who would receive the letters, we can be confident that baseline data are complete -- the IDR includes all prescribing reported by Medicare Advantage plans and stand alone Prescription Drug Plans to CMS.

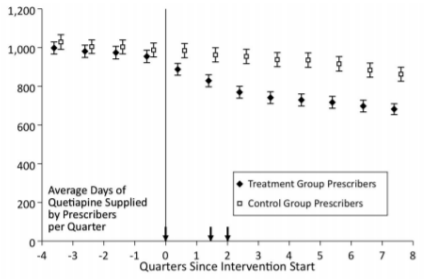
In the event that a prescriber does not show up in the post-treatment data, this will imply that the prescriber changed specialty or state, stopped seeing Medicare patients, retired, or died. For these prescribers, we recommend inputting zeroes for missing years because we assume the true value is zero if the prescriber stopped prescribing. As such, this treatment effect will also capture the effect of prescribers exiting the market.

Descriptive Statistics, Tables, & Graphs

We recommend that a line graph with the primary outcome on the y axis and the month on the x axis, with lines delineating treatment and control, serve as the key figure. The figure included below from a prior collaboration between OES and CMS on the effects of letters sent to potential inappropriate prescribers of quetiapine serves as an example.[4]

[4] Office of Evaluation Sciences. "Reducing Overprescribing in Medicare Part D." 2018.

Figure 1. Quarterly Quetiapine Prescribing in Treatment and Control Groups



** Each point is the average days of quetiapine supplied by prescribers in the quarter. Error bars indicate 95% confidence intervals. Arrowheads denote when letters were sent to prescribers.*

We recommend also including two tables in the analysis:

1. The first table would report pre-treatment prescriber characteristics (baseline values of our primary outcomes, prescriber specialty, and prescriber demographic characteristics as available, as shown in the Table 1 above) by whether or not the prescriber received a letter.
2. The second table would report the effects of the letter on our primary outcomes.
3. An additional Appendix table may report any treatment effects on secondary outcomes (since this would be an exploratory analysis).

Statistical Models & Hypothesis Tests

This section describes the statistical models and hypothesis tests that we recommend to make up the analysis — including any follow-ups on effects in the main statistical model and any exploratory analyses that can be anticipated prior to analysis.

Statistical Models:

We recommend the use of a **difference-in-differences (DD)** design to identify the effect of the peer comparison letters. A simple pre-post design cannot be used to draw causal conclusions because there may be factors other than the receipt of the letter that influence co-prescribing behavior. In order to get around this issue of selection bias and differences in unobservable characteristics between the two groups of prescribers, we recommend using this design that consists of (i) observing the change in prescribing behavior over the pre- to post-intervention period among prescribers in the top 10 percent in terms of benzodiazepine and opioid co-prescribing (2020 letters), and then (ii) comparing this change to the change over the same period observed among untreated prescribers below the top 10 percent cutoff. The80th–90th percentile prescribers will serve as the comparison population (Note that these percentile cutoffs will differ across states and specialties). This design relies on the assumption that absent the peer comparison letters, both groups of prescribers would have the same rates of co-prescribing in the pre-letter and post-letter period (commonly referred to as the "parallel trends assumption").

We recommend relying on the following regression specification for the analysis, which provides an estimate of the impact of prescriber outlier status on the outcome of interest and includes

covariates. The specifications would take the form below (and incorporate the use of heteroskedasticity-robust standard errors):

Specification 1:

$y_{ij} = \beta_0 + \beta_1 month_j * Outlier_i + \beta_2 month_j + \beta_3 Outlier_i + \lambda_i + \sigma_i + \epsilon_{ij}$

where i indexes providers and j indexes month and:

- y_{ij} is the outcome of interest for provider i at month j, i.e. total MME
- β_1 is the causal estimate of interest and represents how the outcome changes among outlier prescribers in that month
- β_2 is a vector for each month in the dataset, where the month the letter was sent out is the referent category
- β_3 is an indicator for whether the prescriber is an outlier, i.e. was above the 90th percentile in the outcome for that specialty and state combination, and 0 otherwise i.e. between 80-90th
- λ_i and σ_i denote an indicator for whether the prescriber is a General Care Prescriber and fixed effects for states, respectively (to control for underlying differences between providers that may affect receiving a letter)
- ϵ_{ij} is the error term

As an additional robustness check, the following regression specification may also be run, which looks at the effects of the letter across the full year after the letter was sent:

Specification 2:

$y_{ij} = \beta_0 + \beta_1 Post_j * Outlier_i + \beta_2 Post_j + \beta_{3yi} Outlier_i + \lambda_i + \sigma_i + \epsilon_{ij}$

where i indexes providers and j indexes all post-treatment periods and:

- y_{ij} is the outcome of interest for provider i in the post-treatment period j, i.e. total MME
- β_1 is the causal estimate of interest and represents how the outcome changes among outlier prescribers in the post-treatment period
- β_2 is an indicator equal to 1 if the data are from the period after letters were mailed, and 0 otherwise
- β_3 is an indicator for whether the prescriber is an outlier, i.e. was above the 90th percentile in the outcome for that specialty and state combination, and 0 otherwise i.e. between 80-90th
- λ_i and σ_i denote an indicator for whether the prescriber is a General Care Prescriber and fixed effects for states, respectively (to control for underlying differences between providers that may affect receiving a letter)
- ϵ_{ij} is the error term

Follow-Up Analyses:

There is a possibility that other factors --- variables not accounted for in the analysis --- might influence observations pre and post intervention. Such factors would reduce our confidence in the DD estimates. Therefore, some additional kinds of analyses may be warranted to check the robustness of the results. However, these analyses are quite time-intensive to run for practical reasons (the OES team standards for statistical software are not available in the CMS virtual data environments but we recommend running these analyses should they be feasible to do so for the CMS team). The Appendix section of this plan describes additional analyses that could be run if the DD results look promising and the data environment allows for more involved specifications. In particular, if there is a directionally negative effect on our main outcome of interest with $p < 0.15$, then the analyses outlined in the Appendix may be valuable to explore.

Additional analyses could also include exploratory outcomes for the 2020 letters: (i) percent of beneficiaries co-prescribed opioids and benzodiazepines, (ii) Diazepam Milligram Equivalents (DME) among co-prescriptions, (iii) opioid and/or benzodiazepine days supplied, and (iv) number of opioid and/or benzodiazepine prescription fills.

Inference Criteria, Including Any Adjustments for Multiple Comparisons:
We recommend the use of a p-value of 0.05 to determine statistical significance (with asterisks according to $^{+}p < .10$, $^{*}p < .05$, and $^{**}p < .01$). All tests should be two-tailed.

Limitations:

Given the relatively small sample size, we expect some limitations in terms of statistical power and thus the analysis may be less equipped to detect smaller effect sizes with precision. As such, we recommend thinking carefully about the implications of confidence intervals generated by results and how this information might inform subsequent, larger-scale trials.

Appendix

Data

The variable names are denoted per their column and table names below. The tables can be found in the IDR, either in the table CLM_LINE in the AAL dataset, or via the ADM platform where CPI has uploaded external tables:

- Quantity of (co-prescribed) opioids dispensed
    - column: CLM_LINE_NDC_QTY
    - table: CLM_LINE
    - Note that this may be computed separately for each week-long period for opioids and benzodiazepines, respectively, for the 2020-22 letters
- Days supply (or prescription length)
    - column: CLM_LINE_DAYS_SUPLY_QTY
    - table: CLM_LINE
- Prescription date (allowing us to identify co-prescribed opioids and benzodiazepines and explore whether the effects of letters decay over time i.e. measure persistence effects)
    - column: CLM_LINE_FROM_DT
    - table: CLM_LINE
- Number of prescription fills
    - column: CLM_LINE_RX_FILL_NUM
    - table: CLM_LINE
- Conversion factors for calculating MME
    - column: CDC_OPIOID_NDC_MME, CDC_MME_STNGTH_PER_UNIT,
    - table: CDC_MME_CF

Other variables may be drawn from the PDE data as well and will include:

- Prescriber specialty and state
    - column: State, Specialty, 80th and 90th percentile prescribing cutoffs, indicator for whether a prescriber was above the 90th percentile cutoff or between the 80th and 90th percentile cutoffs based on baseline data

table: l6jk_DSTNC_NPI_6065 and l6jk_DSTNC_OUTLR_NPI (in ADM)

We recommend calculating the MME (equivalent to the average daily morphine equivalent dose or MED below) based on the following calculations per the Department of Health & Human Services Office of Inspector General "Toolkit: Using Data Analysis To Calculate Opioid Levels and Identify Patients At Risk of Misuse or Overdose," available here. The relevant calculations for the variables are copied below.

$$MED = \frac{(Strength\ per\ unit) \times (Quantity\ dispensed) \times (MME\ conversion\ factor)}{(Days\ supply)}$$

$$Average\ daily\ MED = \frac{Total\ MED\ of\ all\ prescriptions\ in\ the\ timeframe}{Total\ number\ of\ days\ in\ the\ timeframe}$$

Robustness analyses

As such, we recommend only running additional analyses if the DD results look promising. In particular, if a directionally negative effect on the main outcome of interest with $p < 0.15$ is observed, we recommend conducting the two additional analyses described below.

Robustness 1:
If a $p < 0.15$ on the main outcome of interest as per the condition noted above[5] is observed, then we recommend the use of a **synthetic controls** approach as well. A synthetic controls design relies on comparing treated prescribers to a weighted combination of matched control prescribers using data from the pre-treatment period. The sample of control prescribers may be identified based on outcome predictors. In this case, the post-treatment difference (i.e. after the receipt of letters), adjusting for pre-treatment differences, gives the difference-in-difference treatment effects.

Robustness 2:
As an alternative to the DD and synthetic controls approach, if a $p < 0.15$ is observed, as noted above then it is recommended to also use a **regression discontinuity (RD)** design. In the absence of random assignment, this design can compare outcomes for prescribers just above and below the continuous, semi-arbitrary treatment cutoff or threshold wherein letters were sent only to the top 10 percent of prescribers in terms of opioid prescribing (May 2019 letters) or benzodiazepine and opioid co-prescribing (January 2020 letters). This would assume that the probability of being just above or below this cutoff would be random and prescribers on either side of this cutoff are similar in terms of observable and unobservable characteristics. Given that it is essential to pick a bandwidth around the cutoff that is wide enough to have sufficient observations (i.e. power) but not too wide to result in observations that are not comparable, we recommend the use of an optimal bandwidth choice approach (discussed in Cattaneo and Vazquez-Bare, 2016). In this case, the RD design would allow for unbiased estimation of the local average treatment effect (i.e. treatment on treated effect among prescribers near this cutoff) which may have different policy implications since it is simply the effect of those on this margin receiving these letters.

We recommend the DD approach, as feasible, since RD has a particular risk of being underpowered in this setting and because of potentially less precise estimates and lack of external validity when extrapolating results to prescribers further away from the cutoff . The DD approach also has the advantage of being relatively straightforward to communicate to agency partners and other key stakeholders.

[5] We set this $p<0.15$ rule so as to be more inclusive in our results